Assessment of Updated Fuel-Based Emissions Inventories Over the Contiguous United States Using TROPOMI NO2 Retrievals

Li, M.; McDonald, B. C.; McKeen, S. A.; Eskes, H.; Levelt, P.; Francoeur, C.; Harkins, C.; He, J.; Barth, M.; More Authors

DOI
10.1029/2021JD035484

Publication date
2021

Document Version
Final published version

Published in
Journal of Geophysical Research: Atmospheres

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.
Assessment of Updated Fuel-Based Emissions Inventories Over the Contiguous United States Using TROPOMI NO₂ Retrievals

M. Li1,2, B. C. McDonald3, S. A. McKeen1,2, H. Eskes3, P. Levelt1,4,5, C. Francoeur1,2,6, C. Harkins1,2, J. He1,2, M. Barth5, D. K. Henze6, M. M. Bela1,2, M. Trainer1,2, J. A. de Gouw1,7, and G. J. Frost2

1Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA, 2Chemical Sciences Laboratory, NOAA Earth System Research Laboratories, Boulder, CO, USA, 3Royal Netherlands Meteorological Institute, De Bilt, The Netherlands, 4Department of Geoscience and Remote Sensing, Delft University of Technology (TU Delft), Delft, The Netherlands, 5Atmospheric Chemistry Observation & Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, 6Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA, 7Department of Chemistry, University of Colorado, Boulder, CO, USA

Abstract  Nitrogen oxides (NOₓ) are air pollutants critical to ozone and fine particle production in the troposphere. Here, we present fuel-based emission inventories updated to 2018, including for mobile source engines using the Fuel-based Inventory of Vehicle Emissions (FIVEs) and oil and gas production using the Fuel-based Oil and Gas (FOG) inventory. The updated FIVE emissions are now consistent with the NEI17 estimates differing within 2% across the contiguous US (CONUS). Tropospheric NO₂ columns modeled by the Weather Research and Forecasting with Chemistry model (WRF-Chem) are compared with those observed by TROPOspheric Monitoring Instrument (TROPOMI) and Ozone Monitoring Instrument (OMI) during the summer of 2018. Modeled NO₂ columns show strong temporal and spatial correlations with TROPOMI (OMI), identified with biases of −3% (−21%) over CONUS, and +8% (−6%) over point sources plus urban regions. Taking account of the negative bias (~20%) in early version of TROPOMI over polluted regions, WRF-Chem shows good performance with updated FIVE and FOG emissions. Our model tends to under-predict the tropospheric NO₂ columns over background and rural regions (bias of −21% to −3%). Through model sensitivity analyses, we demonstrate the important roles of emissions from soils (11.7% average over CONUS), oil and gas production (4.1%), wildfires (10.6%), and lightning (2.3%) with greater contributions at regional scales. This work provides a roadmap for satellite-based evaluations for emission updates from various sources.

Plain Language Summary  Satellite observations of tropospheric NO₂ columns provide important constraints on air pollutants from space, which have been widely used to validate the performance of atmospheric models. To gain better knowledge of the accuracy of the recently updated fuel-based emissions inventory, we conducted NO₂ assessments between a regional chemical transport model (Weather Research and Forecasting with Chemistry model, WRF-Chem), with the TROPOspheric Monitoring Instrument (TROPOMI) and Ozone Monitoring Instrument (OMI) over the contiguous United States. We find that model simulation results show strong spatial and temporal correlations with satellite observations across point sources, urban, oil and gas production, and rural regions. With updated emissions, our regional atmospheric model can reconcile with satellite retrievals differing from −3% (TROPOMI) to −21% (OMI) overall. Soils, oil and gas production, wildfires and lightning emissions can play key roles in regional air quality. This work provides an important baseline of a pre-COVID year by which sharp changes in anthropogenic NOₓ emissions due to the pandemic can be assessed.

1. Introduction

Nitrogen oxides (NOₓ, NO + NO₂) are important precursors of tropospheric ozone (O₃) and secondary nitrate aerosols, affecting air quality, human health, and climate (Grewe et al., 2012; Hesterberg et al., 2009; Jacob, 1999; Seinfeld & Pandis, 2006; Stocker et al., 2013). Regulated under the Clean Air Act and its Amendments, surface O₃ and aerosol concentrations have decreased in the last decade, although some US cities can still violate the National Ambient Air Quality Standards (NAAQS) set by the Environmental Protection Agency (EPA) (Cooper et al., 2015; Lin et al., 2012; McClure & Jaffe, 2018). Several studies evaluated chemical transport models (CTMs) with atmospheric field observations highlighting potential uncertainties with emissions of NOₓ, a key
precursor to tropospheric ozone formation (Brown-Steiner et al., 2015; Canty et al., 2015; Lin et al., 2012; McDonald et al., 2018; Travis et al., 2016). Quantifying NO$_2$ emissions and resolving their spatial and sectoral distributions are essential for atmospheric modeling and design of ozone and particulate matter control measures from city to regional scales (Duncan et al., 2013; Frost et al., 2006; Jiang et al., 2018; McDonald et al., 2018; Stillman, 1999).

NO$_2$ is detectable from space and is currently retrieved with wide spatial coverage at a relatively high spatial resolution, providing important top-down observational constraints on emissions from various sources (Boersma et al., 2008; Duncan et al., 2013; Kim et al., 2006; Lamsal et al., 2015; Russell et al., 2012; Vinken et al., 2014a, 2014b). Satellite NO$_2$ retrievals with the highest spatial resolution before 2018 are from the Ozone Monitoring Instrument (OMI) onboard the Aura satellite, with a spatial resolution of 13 × 24 km$^2$ at nadir (Levitt et al., 2006, 2018). Top-down inversion estimates of NO$_2$ using regional CTMs and OMI retrievals have been developed, such as Daily Emission estimates Constrained by Satellite Observations (DESCO) using an extended Kalman filter (Ding et al., 2018; Ding et al., 2017; Mijling & van der A, 2012), and multi-year inversions using a hybrid 4D-Var/mass balance inversion method (Qu et al., 2017). Forward modeling is another straight-forward technique to provide top-down evaluation of a-priori emissions based on the remote sensing NO$_2$ columns (Cooper et al., 2017; Lamsal et al., 2010, 2011; Li et al., 2018; Martin et al., 2003). Previous work have pointed out the importance of applying high temporal and spatial resolution NO$_2$ vertical profiles to produce satellite retrievals when comparing to regional models (Lamsal et al., 2010; Laughner et al., 2016). As the successor of OMI, the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor mission was launched in October 2017 (Veeckind et al., 2012). Global NO$_2$ concentration maps are retrieved daily with a higher spatial resolution (3.5 × 7 km$^2$ at nadir, 3.5 × 5.6 km$^2$ since August 6, 2019) and a high signal-to-noise ratio compared to the OMI retrieved data (Beirle et al., 2019; Ding et al., 2020; Goldberg et al., 2019; van Geffen et al., 2020; Veeckind et al., 2012). Recent studies have demonstrated the enhanced capabilities of TROPOMI to provide top-down constraints on NO$_2$ emissions and lifetime over power plants and cities (Beirle et al., 2019; Goldberg et al., 2019; Lorente et al., 2019), even to detect the emission changes due to the COVID-19 pandemic's societal disruption (Ding et al., 2020; Gkatzelis et al., 2021; Liu et al., 2020). Given the plethora of studies that have been published on the air quality impacts during the COVID-19 pandemic (Gkatzelis et al., 2021), including for changes in NO$_2$ emissions, it is even more critical to understand biases in model-satellite NO$_2$ comparisons that existed before the pandemic.

According to the most recent estimates from the US EPA, ~55% of the nation's NO$_2$ emissions are contributed by mobile sources (33% from highway vehicles, 22% from off-highway engines), ~12% by power plants, and ~33% by industry and other area-wide sources (EPA, 2020). These estimates have significant uncertainties, as indicated by previous work based on ground, airborne and satellite observations, especially for agricultural, industrial, and off-road sources (Anderson et al., 2014; Jiang et al., 2018; McDonald et al., 2018; Souri et al., 2016). Satellite data provide essential top-down evaluations to quantify the impacts from emissions of various sectors (Jiang et al., 2018; Liu et al., 2017; Sha et al., 2021). Jiang et al. (2018) suggested discrepancies in the trend of mobile source emissions when compared with ground and satellite observations across the contiguous United States (CONUS). The NO$_2$ lifetime has also changed as a result of reduced NO$_2$ emissions and nonlinear chemistry over US cities (Laughner & Cohen, 2019). On the other hand, due to the emission reductions from mobile sources and power plants, the free tropospheric NO$_2$ background is becoming relatively more important, according to recent satellite evaluations (Qu et al., 2021; Silvern et al., 2018, 2019).

Since there is often a lag of a few years in the availability of emission inventories, an up-to-date evaluation of US bottom-up NO$_2$ inventories is challenging to create. For example, the National Emissions Inventory (NEI) 2017 was released in 2020, and detailed analyses of the inventory are still forthcoming (EPA, 2020). To support NOAA air quality forecasting efforts, the Fuel-based Inventory of Vehicle Emissions (FIVEs) have been developed for mobile source engines and implemented in the Weather Research and Forecasting with Chemistry (WRF-Chem) model (Kim et al., 2016; McDonald et al., 2018) and has been implemented in the Rapid Refresh with Chemistry (RAP-Chem) forecast model (https://rapidrefresh.noaa.gov/RApChem/). Recently, the FIVE inventory has been updated through the end of 2020 and captures changes through the COVID-19 pandemic (Harkins et al., 2021).

In this study, we focus on evaluations of FIVE with TROPOMI NO$_2$ data from 2018, leveraging prior modeling by the NOAA Chemical Sciences Laboratory performed for the Long Island Sound Tropospheric Ozone Study (LISTOS) (Cogggon et al., 2021). Here, we extend the analysis to CONUS and evaluate urban-rural distributions
in NO$_2$ columns. In rural regions, we also assess NO$_2$ emissions from oil and gas development activities, along with other anthropogenic (e.g., agricultural soils) and natural (e.g., lightning) sources. The objectives of this study are to: (a) reconcile US anthropogenic NO$_2$ emissions with NO$_2$ retrievals of TROPOMI through the use of a regional 3-D CTM; and (b) provide insights about how further development in biogenic and anthropogenic emission inventories could improve the reconciliation of NOx inventories with remote sensing tropospheric NO$_2$ columns. We relied on the construction of bottom-up inventories in conjunction with forward chemical transport modeling to address these objectives. Finally, through a systematic evaluation of a pre-COVID-19 years, we aim to establish a baseline by which to evaluate changes in satellite NO$_2$ columns due to the COVID-19 pandemic in the future.

2. Materials, Methods, and Data

2.1. WRF-Chem Model Setup

The Weather Research and Forecasting with Chemistry (WRF-Chem) (version 4.0) model was set up to simulate air quality (Grell et al., 2005) at a horizontal spatial resolution of 12 × 12 km$^2$ over the CONUS during summer (June–August) in 2018. The modeled NO$_2$ concentrations between 13:00 and 14:00 local time for each grid cell are extracted and averaged to make consistent comparisons with TROPOMI. The model configurations in this study are summarized in Table S1 in Supporting Information S1. Vertically, the model extends from the surface to 50 hPa, with the intervening atmosphere divided into 50 layers. We use the Grell-Devenyi (GD) ensemble cumulus scheme in the WRF model. The model chemistry was based on the RACM_ESRL gas phase chemistry mechanism (Stockwell et al., 1997), with updated isoprene reaction rates (Kim et al., 2009). The North American Mesoscale Model data was used to provide initial and boundary conditions for the meteorology in the model. We applied the default lightning scheme based on the convective cloud-top height in WRF-Chem to simulate contributions from lightning NO$_2$ (Wong et al., 2013). Anthropogenic and other natural emissions provided to the model are described below.

2.1.1. Anthropogenic Emissions

We updated the anthropogenic emissions inventory to 2018 levels by:

1. Utilizing the FIVEs for mobile sources including both on-road and off-road engines (McDonald et al., 2014, 2018). The inventory utilizes fuel sales records to quantify mobile source activity, and NO$_2$ emission factors normalized to fuel use based on near-roadway and tunnel study measurements and laboratory studies. More details on the development of FIVE can be found in previous publications (McDonald et al., 2014). The FIVE inventory has been evaluated with ground and aircraft-based measurements at both regional (Kim et al., 2016) and continental scales (McDonald et al., 2018). FIVE was shown to improve model predictions of NO$_2$ and O$_3$ when compared to the NEI (McDonald et al., 2014). The FIVE inventory has been updated from 2018 through the end of 2020, including the period of the COVID-19 pandemic (Harkins et al., 2021). Here, we use the FIVE emissions for 2018 (denoted FIVE18) in the baseline simulations.

2. Utilizing the Fuel-based Oil and Gas (FOG) Inventory for oil and natural gas production regions over CONUS (Francoeur et al., 2021). According to the top-down evaluations using aircraft observations, over basins accounting for ∼25% of US onshore oil and gas production, oil and gas NO$_2$ emissions from the FOG inventory are within ∼10% of the aircraft-derived emissions (Francoeur et al., 2021).


4. For other point and area sources (e.g., large industrial facilities), emissions are from the NEI 2014 or 2017. Most of the sensitivity runs were performed using the NEI 2014. We added a baseline simulation incorporating the NEI17 to reflect the most recent emission updates and ensure our results did not change significantly. It is been shown that relative to power plants and mobile sources, industrial and area source emissions have not changed much in recent years (Jiang et al., 2018).

5. Canada and Mexico emissions are potentially important for background sources to the US. They are not included in NEI 17 or NEI 14, and were initially taken from defaults provided by the WRF-Chem community, which were based on the NEI 2005 and emissions scaled to more recent years based on the Community Emissions Data System (CEDS) (Hoesly et al., 2018). We also perform the baseline simulation updating Canada and Mexico emissions using the Copernicus Atmosphere Monitoring Service (CAMS) inventory (Granier
et al., 2019), which is gridded at 0.1° × 0.1° globally. We include the shipping emissions from CAMS, and exclude these emissions from the NEI 2017 to avoid double-counting.

We define two baseline emission cases summarized in Table 1. The first baseline case (Baseline I) is defined as FIVE18 (mobile source) + CEMS (power plant) + FOG (oil & gas) + NEI17 (other area/point) + CAMS (CAN/MEX/Ocean) emissions. Baseline I is set up to reflect that the combination of NEI17 for other area and point source emissions and the FIVE18 for mobile sources. This baseline represents the most up-to-date emissions available for our period of interest. To analyze the model performance by sectors, we group the grid cells according to the locations of point sources and cities, and conduct sensitivity simulations to identify the dominant background sources (soil, oil and gas production, wildfire, and lightning). Since the NEI17 case was released in the past year, many of the sensitivity runs were performed previously using the NEI14. We rely on Baseline II, which utilizes the NEI14 for other area and point source emissions instead of NEI17, to assess relative changes in tropospheric NO$_2$ columns due to individual emission sectors.

### Baseline and Sensitivity Model Cases Set Up in WRF-Chem

<table>
<thead>
<tr>
<th>Model cases</th>
<th>Mobile</th>
<th>Biogenic</th>
<th>Point</th>
<th>Other</th>
<th>O&amp;G</th>
<th>Fire</th>
<th>Lightning</th>
<th>Canada/Mexico</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline I</td>
<td>FIVE18</td>
<td>BEIS</td>
<td>CEMS</td>
<td>NEI17</td>
<td>FOG</td>
<td>n/a</td>
<td>n/a</td>
<td>CAMS</td>
</tr>
<tr>
<td>Baseline II</td>
<td>FIVE18</td>
<td>BEIS</td>
<td>CEMS</td>
<td>NEI14</td>
<td>FOG</td>
<td>n/a</td>
<td>n/a</td>
<td>NEI05$^b$</td>
</tr>
<tr>
<td>Soil</td>
<td>FIVE18</td>
<td>BEIS, excluding soil NOx</td>
<td>CEMS</td>
<td>NEI14</td>
<td>FOG</td>
<td>n/a</td>
<td>n/a</td>
<td>NEI05$^b$</td>
</tr>
<tr>
<td>Oil &amp; Gas</td>
<td>FIVE18</td>
<td>BEIS</td>
<td>CEMS</td>
<td>NEI14</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>NEI05$^b$</td>
</tr>
<tr>
<td>Fire</td>
<td>FIVE18</td>
<td>BEIS</td>
<td>CEMS</td>
<td>NEI14</td>
<td>FOG</td>
<td>3BEM$^a$</td>
<td>n/a</td>
<td>NEI05$^b$</td>
</tr>
<tr>
<td>Lightning</td>
<td>FIVE18</td>
<td>BEIS</td>
<td>CEMS</td>
<td>NEI14</td>
<td>FOG</td>
<td>n/a</td>
<td>Wong et al. (2013)</td>
<td>NEI05$^b$</td>
</tr>
</tbody>
</table>

$^a$The Brazilian Biomass Burning Emission Model (3BEM) (Longo et al., 2010). $^b$NEI05 emissions scaled to 2014 using trends for Canada and Mexico from the Community Emissions Database System (CEDS) inventory (Hoesly et al., 2018).

We define two baseline emission cases summarized in Table 1. The first baseline case (Baseline I) is defined as FIVE18 (mobile source) + CEMS (power plant) + FOG (oil & gas) + NEI17 (other area/point) + CAMS (CAN/MEX/Ocean) emissions. Baseline I is set up to reflect that the combination of NEI17 for other area and point source emissions and the FIVE18 for mobile sources. This baseline represents the most up-to-date emissions available for our period of interest. To analyze the model performance by sectors, we group the grid cells according to the locations of point sources and cities, and conduct sensitivity simulations to identify the dominant background sources (soil, oil and gas production, wildfire, and lightning). Since the NEI17 case was released in the past year, many of the sensitivity runs were performed previously using the NEI14. We rely on Baseline II, which utilizes the NEI14 for other area and point source emissions instead of NEI17, to assess relative changes in tropospheric NO$_2$ columns due to individual emission sectors.

#### 2.1.2. Biogenic Emissions

For biogenic emissions, we utilized the Biogenic Emissions Inventory System (BEIS3.13) model with predicted isoprene emissions reduced by a factor of two (Pierce et al., 1998, 2002). An adjustment of isoprene emissions was performed to reconcile the modeled HCHO columns as well as chemistry with TROPOMI observations. Figure S1 in Supporting Information S1 demonstrates that with the 50% isoprene scaling factor applied, the model-TROPOMI correlations are significantly improved over the Southeastern US (bias was improved from +40% to +8%).

BEIS3.13 estimates the biogenic NO$_2$ emissions that result from microbial activity from soil in the growing season (April–October) and non-growing season (November–March) for both agricultural and non-agricultural areas. BEIS3.13 resolves the forest canopy coverage by tree species based on a 1-km resolution vegetation database. For the summer growing season in our study, the model applies a soil NO$_2$ emission estimation algorithm that accounts for soil moisture, crop canopy coverage, and fertilizer application (Guenther et al., 2000; Pouliot & Pierce, 2009; Vukovich & Pierce, 2015). The soil emissions are calculated in-line with the meteorology within the CTM.

#### 2.1.3. Wildfire Emissions

The wildfire emissions are estimated from the Brazilian Biomass Burning Emission Model (3BEM), which combines remote-sensing fire products and field observations (Longo et al., 2010). Following the approach of Freitas et al. (2005), 3BEM estimates the mass of targeted species by multiplying the amount of above-ground biomass by the combustion factor, the emission factor, and the burnt area (Freitas et al., 2005; Longo et al., 2010). The burnt area estimates are based on three fire products from satellite retrievals using a hybrid algorithm or statistical analyses of the fire scars. Emissions of chemical tracers are calculated based on the vegetation-related fire properties from field observations. Top-down evaluations using MOPITT data suggest the 3BEM model emission estimates are reliable at regional scales (Longo et al., 2010). 3BEM is the default fire package for WRF-Chem
with RACM chemistry, and it’s under development according to the information derived from field measurements in the US (Grell et al., 2011). Assessing various fire emission inventories is important but beyond the scope of this paper. In this work, we applied 3BEM in the sensitivity case simulations to evaluate the contributions to NO$_2$ columns from fires.

2.1.4. Lightning NO$_x$

We applied the lightning parameterization in WRF-Chem based on cloud-top height ($z_{\text{top}}$) (Price and Rind, 1992; Wong et al., 2013). These relationships between lightning flash rate and cloud-top height were established via an empirical framework combined with observational data (Williams, 1985). The equations of lightning flash rate ($f_c$) over continental and marine clouds are:

\begin{align}
    f_c &= 3.44 \times 10^{-5} \frac{z_{\text{top}}}{1.73} \quad \text{(continental)} \\
    f_c &= 6.2 \times 10^{-4} \frac{z_{\text{top}}}{4.73} \quad \text{(marine)}
\end{align}

By using a 12-km grid spacing in our simulations, the Grell-Devenyi ensemble convection parameterization scheme computes the cloud top height using the level of neutral buoyancy (LNB) which is based on the maximum moist static energy. The partitioning ratio between intra-cloud and cloud-to-ground flashes (IC:CG ratio) was determined from the prescribed climatological mean of Boccippio et al. (2001), based on the Optical Transient Detector and the National Lightning Detection Network (Boccippio et al., 2001). We applied a moderate NO$_x$ production rate of 125 moles per flash in the calculation (Pickering et al., 2016; Pollack et al., 2016; Schumann and Huntrieser, 2007). The vertical distribution of the lightning-generated NO$_x$ places most of the NO$_x$ in the upper troposphere (Ott et al., 2010). The detailed parameterizations are listed in Table S1 in Supporting Information S1.

The monthly averaged lightning NO$_x$ emissions are 2,660 metric tons of NO$_2$ per day (6% of the total emissions over the whole domain), concentrating in the Gulf of Mexico, Arizona, and Southeast US, as shown in Figure S2 in Supporting Information S1.

2.2. TROPOMI/OMI NO$_2$ Product

The NO$_2$ column data of TROPOMI is processed by the Royal Netherlands Meteorological Institute (KNMI) and made available from the Sentinel-5P data portal (https://s5phub.copernicus.eu/dhus/#/home, Last access: November 19, 2019). We obtained the daily NO$_2$ level 2 swath data (version 1.2.2) from the Reprocessing stream (RPRO), and processed the data into summer-averaged tropospheric NO$_2$ columns for model evaluation. To reduce retrieval uncertainties, daily data with high quality flag values ($qf \geq 0.75$) are included, which screens out pixels that are covered at least partially by cloud (cloud radiance fraction >0.5), by snow/ice, and those that have problematic retrievals (Eskes et al., 2020).

The tropospheric NO$_2$ columns are sensitive to the NO$_2$ vertical profiles and the calculated air mass factor (AMF). It is important to incorporate the AMF and averaging kernels into the comparison between modeled and TROPOMI columns (Boersma et al., 2004; Eskes & Boersma, 2003a; Lamsal et al., 2010). In this work, we combined the averaging kernels ($A_k$ is the total, $A_k^{\text{trop}}$ is the tropospheric kernel) at each level $k$ for filtered swath pixels with the NO$_2$ vertical profile shapes ($X_k^{wc}$) from the WRF-Chem simulation. This process eliminates the bias in model-satellite comparisons introduced by the a-priori profile assumptions from a global transport model Tracer Model 5 (TM5) (Huijnen et al., 2010). It is shown that such revisions utilizing the NO$_2$ vertical profile shapes derived from WRF-Chem and averaging kernels make the comparisons more consistent (Eskes & Boersma, 2003b; Lamsal et al., 2010). The standard TROPOMI NO$_2$ data were revised based on Equations 3–5 as below (Eskes et al., 2020; Lamsal et al., 2010; Palmer et al., 2001):

\begin{align}
    \Omega_{wc} &= \frac{M^{\text{trop}}_{wc}}{M^{\text{trop}}_{\text{wc}}} \Omega_m \\
    M^{\text{trop}}_{\text{wc}} &= \frac{M^{\text{trop}} \sum_k A_k^{\text{trop}} X_k^{wc}}{\sum_k X_k^{wc}} \\
    A_k^{\text{trop}} &= \frac{M}{M^{\text{trop}}} A_k
\end{align}
where $h$ is the tropopause, $\Omega_{\text{wc}}/\Omega_{\text{tm}}$ denote the revised/standard satellite tropospheric vertical NO$_2$ column density (in molecules cm$^{-2}$), $M/M^{\text{a-priori}}$ are the total/tropospheric AMF derived from the a-priori TM5 model, and $M^{\text{a-priori}}$ is the derived tropospheric AMF with WRF-Chem simulated vertical profile shape ($X^w_{\text{a-priori}}$). Updating the AMF using the higher-resolution WRF-Chem model results in 20% higher concentrations on average over urban areas, and 3% decreases over lower concentration areas in background and rural regions. These significant effects further demonstrate the importance of AMF revision based on high resolution regional models and averaging kernels. All available swath daily data were resampled based on a “conservative” algorithm (https://xesmf.readthedocs.io) to model grids at a resolution of 12 × 12 km$^2$ to match the model resolution, and further averaged for the period of the summer of 2018. By averaging the TROPOMI data over a long time period (3 months in our study) and applying the AMF correction based on NO$_2$ profiles from our regional model, we assume that the random errors due to instrumental noise are partly eliminated.

The systematic error in the TROPOMI NO$_2$ tropospheric column is mainly related to the slant column spectral fitting, separation of stratospheric and tropospheric columns, and uncertainties in model parameters to derive AMFs (Veefkind et al., 2012). Validation using ground-based Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) NO$_2$ measurements shows that TROPOMI tropospheric column data are biased low by −23% to −37% in slightly polluted regions, and up to −51% in highly polluted conditions (Verhoelst et al., 2021). Compared to the airborne and ground-based measurements of PANDORA (NASA Goddard Space Flight Center) (Herman et al., 2009) for the 2018 Long Island Sound Tropospheric Ozone Study (LISTOS), TROPOMI tropospheric NO$_2$ columns show an overall negative bias of 19%–33% (Judd et al., 2020). The bias estimate is reduced to −7% to −19% with updated profiles from regional models (Judd et al., 2020). Similarly, Wang et al. (2020) found a low bias of 30%–50% in the TROPOMI trop. NO$_2$ columns as evaluated based on ground MAX-DOAS instruments in China. Identified causes of the systematic low bias are the surface albedo (the TROPOMI retrieval still uses the OMI Lambert Equivalent Reflectance data, with a suspected high bias), the cloud retrieval (and implicit aerosol correction) and the profile shapes. Surface albedo uncertainties can lead to 20% uncertainties of NO$_2$ locally and systematic biases regionally. In TROPOMI versions 1.2.x and 1.3.x the cloud pressure was found to be overestimated, leading to a low NO$_2$ bias of the order of 12% over the US compared to the OMI QA4ECV product in winter, and smaller biases in summer. The a-priori profile shape used in the retrieval strongly influences the direct comparisons with MAX-DOAS, PANDORA, and CTMs, but does not influence comparisons with models when the averaging kernels are used in this study. Accounting for this, we still expect in general a low bias up to −20% in the TROPOMI tropospheric NO$_2$ data over polluted regions depending on locations and seasons. We further conduct similar ground-based evaluations of the TROPOMI tropospheric NO$_2$ data with AMF correction using the PANDORA measurements over New York City (shown in Figure S3 in Supporting Information S1). A systematically 20% low bias is found in TROPOMO over polluted cities, consistent with the findings in previous satellite validation work (Lambert et al., 2021).

The sum of stratospheric and upper tropospheric NO$_2$ columns is kept consistent with the TROPOMI observations through the assimilation step in the retrieval, but the separation of the column into a stratospheric and tropospheric part is based on the modeled profiles of the TM5-MP model. A bias between the free tropospheric column in TM5-MP and WRF-Chem will directly manifest itself in biases over rural background regions with low tropospheric column amounts. We estimate the seasonal averaged error of TROPOMI NO$_2$ data by propagation of errors with the averaging kernels applied in each swath (as shown in Figure S4 in Supporting Information S1). The spatial pattern indicate that the standard deviations of satellite retrievals can be large over remote regions with low concentrations, and averaging over a long period of time is especially important to interpret these data. We filter out those grids with very large uncertainties (standard deviation >100%) in our statistical analyses.

OMI provides an independent satellite retrieval of NO$_2$ by which to assess uncertainties in TROPOMI NO$_2$ in the model comparison and evaluation. The OMI tropospheric NO$_2$ column V003 Level 2 swath data at NASA Goddard Earth Science (GES) Data and Information Services (DISC) were obtained from the NASA’s Earth Observing System Data and Information System (EOSDIS) (https://earthdata.nasa.gov/). Last access: January 14, 2021) for June–August in 2018 (Krotkov et al., 2019), and then further processed into seasonal averages. We screened out pixels of OMI following the principles of TROPOMI filtering. Daily pixels with cloud fraction >30% (in accordance with cloud radiance fraction >50%), solar zenith angle >85°, or those affected by the row anomaly are excluded before averaging following the user manual guide (https://aura.gesdisc.eosdis.nasa.gov/data/Aura_OMI_Level2/OMNO2.003/doc/README.OMNO2.pdf). Only grids with more than 10 valid observations
during the summer (3 months) are included in the statistical analyses. Compared to the ground measurements, OMI standard tropospheric NO$_2$ product tend to be biased low by up to 50% over polluted regions (Compernolle et al., 2020; Lamsal et al., 2014; Wang et al., 2020), and biased high in remote areas (Lamsal et al., 2014). To exclude the effect of the a-priori vertical profiles from global models, the OMI NO$_2$ standard product data also follow the same AMF revision procedure as above to compare with WRF-Chem in the assessment.

3. Results and Discussions

3.1. Model Evaluations Over the CONUS

The sectoral emissions used in this study are summarized in Table 2. In contrast to McDonald et al. (2018), which suggested that mobile source (on-road + off-road) NO$_x$ emissions in the NEI 2011–14 were ~30% higher than FIVE in 2013, we note that the emission estimates in the most recent years are in better agreement. As shown in Table 2, the mobile source emissions totals from FIVE18 and NEI17 are within 2%, and differ by 8% for on-road sources. This convergence can be explained by faster decreases in NEI emissions than in FIVE in the intervening years between 2013 and 2018 (Jiang et al., 2018). The areawide emissions increased by 34% from 2014 to 2017 as estimated by NEI, which may reflect differences in methodology in how these emissions are processed. Soil, lightning, and Canada/Mexico/ocean-going vessels NO$_x$ emissions contribute 11%, 6%, and 20% of the total for the whole domain, respectively.

Before evaluating the emissions inventory in the WRF-Chem model and with TROPOMI NO$_2$ columns, we perform a meteorological evaluation with the NOAA Integrated Surface Database (https://www.ncdc.noaa.gov/isd), as summarized in Table 3. The evaluation is conducted over the 9 NOAA climate regions (https://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php). For each region, we plot a time series of the median, 25th and 75th percentiles from the model and with the surface observations for surface temperature, wind speed, and water vapor concentration at the satellite overpass time (see Figure S5 in Supporting Information S1 for example). The bias and $R^2$ are calculated on the median of the model and observations across the time series.

In a broad regional sense, the model performance regarding meteorology is generally good, with biases lying in the range of −0.5% to +0.1% for surface temperature, −0.2% to +15.5% for water vapor density, and −10.8% to +6.3% for wind speed, and quite high correlation coefficients for all meteorological fields ($R^2 = 0.63–0.98$).

Figure 1 compares the tropospheric NO$_2$ columns modeled by WRF-Chem and those from TROPOMI during the summer of 2018 for the Baseline NEI17 emissions case. Comparisons between model and TROPOMI by the US regions (West, Midwest, South, and Northeast) are demonstrated in Figure S6 in Supporting Information S1. Qualitatively, our model and the satellite show generally consistent spatial patterns (see Figures 1a and 1b). Overall, the Baseline NEI17 simulation of NO$_2$ columns agrees well with the TROPOMI NO$_2$ data at a continental scale ($R^2 = 0.78$, slope = 0.93, Theil-Sen linear regression with fixed intercept as zero, NMB = −3%). Enhanced
NO$_2$ columns due to emissions from cities, point sources, and oil and gas basins are seen in both the model and satellite data. However, discrepancies are also found over several hotspot regions (e.g., New York, Chicago), as well as the rural regions over the Eastern US (see Figures 1c and 1d). Emissions, meteorology and chemistry can all contribute to the model-satellite differences at regional scales. Based on the above analyses, it is unlikely that biases in meteorology (maximum of $-10.8\%$ for wind speed) can explain the regional model biases in NO$_2$.

We evaluated different emission sectors in the inventory with TROPOMI, leveraging the high spatial resolution of the satellite retrieval and the fact that various emission sectors exhibit different spatial patterns in the inventory. We grouped grid cells by those dominated by emissions from point, cities and background sources (agriculture, Region

<table>
<thead>
<tr>
<th>Region</th>
<th>$T_2$ ($^\circ$C)</th>
<th>WS (m/s)</th>
<th>$C_{NO_2}$ (g/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td>$-0.3%$, $R^2 = 0.96$</td>
<td>$+7.4%$, $R^2 = 0.84$</td>
<td>$+1.2%$, $R^2 = 0.98$</td>
</tr>
<tr>
<td>Southeast</td>
<td>$+0.0%$, $R^2 = 0.89$</td>
<td>$+12.6%$, $R^2 = 0.82$</td>
<td>$-0.6%$, $R^2 = 0.96$</td>
</tr>
<tr>
<td>Ohio Valley</td>
<td>$+0.0%$, $R^2 = 0.93$</td>
<td>$+12.5%$, $R^2 = 0.92$</td>
<td>$-2.7%$, $R^2 = 0.94$</td>
</tr>
<tr>
<td>Upper Midwest</td>
<td>$+0.1%$, $R^2 = 0.92$</td>
<td>$+15.5%$, $R^2 = 0.90$</td>
<td>$-1.6%$, $R^2 = 0.97$</td>
</tr>
<tr>
<td>South</td>
<td>$+0.0%$, $R^2 = 0.88$</td>
<td>$+2.5%$, $R^2 = 0.83$</td>
<td>$-1.5%$, $R^2 = 0.85$</td>
</tr>
<tr>
<td>Rockies</td>
<td>$+0.0%$, $R^2 = 0.96$</td>
<td>$+1.1%$, $R^2 = 0.88$</td>
<td>$-8.0%$, $R^2 = 0.83$</td>
</tr>
<tr>
<td>Southwest</td>
<td>$-0.5%$, $R^2 = 0.87$</td>
<td>$+7.6%$, $R^2 = 0.82$</td>
<td>$+5.6%$, $R^2 = 0.94$</td>
</tr>
<tr>
<td>Northwest</td>
<td>$-0.4%$, $R^2 = 0.94$</td>
<td>$+6.8%$, $R^2 = 0.79$</td>
<td>$+6.3%$, $R^2 = 0.63$</td>
</tr>
<tr>
<td>West</td>
<td>$-0.3%$, $R^2 = 0.93$</td>
<td>$-0.2%$, $R^2 = 0.74$</td>
<td>$-10.8%$, $R^2 = 0.86$</td>
</tr>
</tbody>
</table>

Figure 1. NO$_2$ tropospheric columns in (a) WRF-Chem simulated with the updated fuel-based inventories for June–August 2018, (b) TROPOMI retrievals with AMF revisions, (c) Absolute differences and (d) Relative differences between WRF-Chem and TROPOMI. The top 10 densely populated cities (solid rectangles) and oil and gas production basins (dotted rectangles) are marked in (c) and (d).
oil and gas production, lightning, and wildfire). Grids were assigned to each source group according to the following procedures. First, if the a-priori emission proportion of point sources/mobile sources is greater than 60% as determined by the bottom-up inventory, then the grid is marked as “point”/“city.” Point sources consist of power plants and industrial sources, and we define point source regions as within 50 km distance of a designated point source. This distance-based limit is derived by assuming NO\(_2\) lifetime is 3 h for point sources and an average wind speed of 5 m s\(^{-1}\) in summer (Liu et al., 2016). In Figure 2b, we compare “point sources” and “city” grids together because often these sources are in close proximity to one another. Some point sources are isolated from cities, and these are shown in panel 2c. Panel 2d shows results for the grid cells over the top 10 urbanized areas in the US (cities listed in Table 4). Grids that were neither affected by “point sources” nor “city” are marked as “background and rural” in panel 2e.

We explore the effect of various linear regression methods on the model-satellite evaluation, including Theil-Sen regression (TSR), Orthogonal distance regression (ODR), and ordinary least square regression (OLS). As shown in Figure S7 in Supporting Information S1, for each statistical method, fixing the intercept or not has minor effects on our analysis. The regression slopes using OLS and ODR can be 13%–29% higher than the TSR's. As discussed in Vigouroux et al. (2020), TSR is recommended for satellite validation because it can give reliable results in the presence of outliers and/or heteroscedasticity (Vigouroux et al., 2020). In the following analyses, we apply TSR with zero-intercept, and NMB as statistical metrics to perform model-satellite evaluations using TROPOMI and OMI data.

As shown in Figure 2a, the overall bias of WRF-Chem over CONUS compared to TROPOMI is −3% (NMB, panel a), +8% for point sources plus cities (panel b), and −3% for background and rural regions (panel c). The good correlation statistics (slope = 0.94, \(R^2 = 0.78\)) demonstrate the consistent spatial distributions for the emissions and model-simulated and the satellite columns. WRF-Chem agrees well with TROPOMI for point and city sources (Figure 2b), within 10%. For background regions, the uncertainties of the satellite data (see Figure S4 in Supporting Information S1), as well as the free tropospheric NO\(_2\) simulations in models (Silvern et al., 2019) are higher than the polluted areas. Figure 2e demonstrates that the performance of model is also good (slope = 0.94, \(R^2 = 0.78\), NMB = −3%) compared to the TROPOMI observations and within the uncertainty of the TROPOMI NO\(_2\) background values.

OMI provides the uncertainty bounds from satellite retrievals along with TROPOMI. The OMI NO\(_2\) columns are higher than TROPOMI observations by 22% overall (see Figure S8 in Supporting Information S1). The ~20% differences between TROPOMI and OMI data are consistent with other validation results (Wang et al., 2020), due to different cloud masks, AMFs, a-priori vertical shapes, row anomalies, spatial resolutions, and so on. While we concentrate our emissions inventory and model evaluation below with TROPOMI NO\(_2\) to different cloud masks, AMFs, a-priori vertical shapes, row anomalies, spatial resolutions, and so on. While we concentrate our emissions inventory and model evaluation below with TROPOMI NO\(_2\), we use OMI as a means to bound the uncertainty in satellite retrievals of NO\(_2\) over different emission source regions (Kim et al., 2009).

As shown in Figure 2, in general, the correlations between WRF-Chem and OMI are not as high as those with TROPOMI. Over CONUS, modeled columns tend to be lower than those observed by OMI by 21% (slope = 0.78, \(R^2 = 0.69\)). Like with TROPOMI, we find good agreement over point and city sources between WRF-Chem and OMI (slope = 0.85, NMB = −6%, Figure 2b). For background regions and oil and gas (O&G) basins, columns modeled by WRF-Chem are ~20% lower than those observed by OMI (see Figures 2e and 2f). In summary, despite the uncertainties in the satellite data inferred by differences in TROPOMI and OMI, modeled tropospheric NO\(_2\) columns can reconcile with satellite data within ~20% over point sources, cities, and background regions. More evaluations based on aircraft or ground measurements of TROPOMI and OMI tropospheric NO\(_2\) data over background regions with low concentrations are needed in future work. Next, we focus our results and discussion on more targeted analysis of emission source regions.

### 3.2. Model Evaluation Over Isolated Point Sources

We show the relative column differences between WRF-Chem and TROPOMI, labeled with the locations of point sources in Figure S9 in Supporting Information S1. It should be noted that the point sources include power plants that have CEMS installed and monitored on an hourly basis, and industrial point sources without online stack monitoring. To exclude the effect of cities, we selected the 565 isolated point sources away from cities (>50 km) and show the correlations in Figure 2c. The systematic bias of isolated point sources across the US is −8% (slope = 0.91, \(R^2 = 0.66\)), which is consistent with recent top-down emission derivations using TROPOMI over point sources using model independent techniques (Goldberg et al., 2019). Below we provide an assessment
Figure 2. Comparison of July–August 2018 averaged tropospheric NO$_2$ columns from WRF-Chem versus TROPOMI/OMI over (a) The whole CONUS domain, (b) Point sources and cities, (c) Isolated point sources, (d) Top 10 cities by population, (e) Background and rural, (e) Oil and gas basins. The criteria for grid selection by emission source are described in the main text of Section 3. We performed the Theil-Sen multivariate linear regressions with intercept fixed as zero. OMI-model comparisons (orange circles) are included to provide bounds for satellite data. Less data points are available for OMI compared to TROPOMI in these comparisons, due to the row anomalies and the relatively coarse spatial resolution in OMI.
of whether this bias is significant based on uncertainties in bottom-up emission inventories, effects of non-linear chemistry, and satellite NO$_2$ observations.

Previous evaluations of power plant emissions from CEMS based on aircraft measurements derived a total 2σ uncertainty of ±24%, and an accuracy of stationary emissions of NO$_x$ within 10% (Frost et al., 2006; Placet et al., 2000), providing an uncertainty limit of CEMS power plant emissions of 10%–24%. Apart from uncertainty in bottom-up emissions, the non-linear chemistry of NO$_2$ is another potential factor contributing to the model biases. Previous work has shown that regional models at a resolution of 4–12 km are sufficient to minimize biases associated with the nonlinear chemistry of NO$_2$ during plume dilution downwind of power plants, within 10% (Valin et al., 2011). Finally, the standard deviations provided by TROPOMI NO$_2$ over these isolated point sources are 17%–97% (median average of 42%), illustrating the large variability in satellite observations (see Figure S4 in Supporting Information S1) and the need to perform model-satellite evaluations over multiple months to enhance signal-to-noise in the satellite NO$_2$ data. Following the method of Kim et al. (2009), we assign 20% uncertainties in the satellite retrievals through comparisons between TROPOMI and OMI tropospheric NO$_2$ data (Figure S8 in Supporting Information S1). Overall, we are able to reconcile isolated point source emissions simulated in a 12 × 12 km$^2$ CONUS WRF-Chem model to within ~20% of satellite TROPOMI and OMI NO$_2$ observations. The model-satellite differences are within the uncertainty of the point source emissions inventory (10%–24%), non-linear chemistry (10%), and satellite NO$_2$ retrievals (±20%) (Kim et al., 2009).

### 3.3. Model Evaluation Over the 10 Most Densely Populated Cities

WRF-Chem NO$_2$ columns simulated with the Baseline I are 27% and 4% higher than the TROPOMI and OMI columns, respectively, over the 10 most populated US cities (see Figure 2d). Previous validation work on TROPOMI using MAX-DOAS and PANDORA ground measurements demonstrate that TROPOMI NO$_2$ columns are biased low by ~50% over polluted urban areas (Verhoeest et al., 2021; Wang et al., 2020). We also found ~20% biases in the current version of TROPOMI NO$_2$ data with AMF correction over cities (as illustrated in Section 2.2), which can largely explain the differences between WRF-Chem and TROPOMI in this comparison. Uncertainties in the bottom-up mobile source NO$_x$ emissions are estimated to be ±15% from FIVE (Jiang et al., 2018). We chose to limit our analysis to the top 10 biggest cities so we can also assess the spatial variability of NO$_2$ across the metropolitan region in the model with TROPOMI data (in each case, there are at least 24 grid cells per city). The statistics for the top 10 cities in terms of population size are summarized in Table 4. The grid numbers are between 24 and 64, and the averaged valid satellite retrievals for each grid are 18–78 over summer for these cities. Thus the numbers of valid retrievals for the top biggest 10 cities are in the range of 432–3,282 over the investigated period (3 months). The large amount of observations provide confidence in the statistical analyses of our study. Los Angeles (LA) NO$_2$ columns are biased +15% in WRF-Chem compared to TROPOMI.

#### Table 4

<table>
<thead>
<tr>
<th>Top 10 cities</th>
<th>WRF-Chem (10$^{15}$ mole cm$^{-2}$)</th>
<th>TROPOMI (10$^{15}$ mole cm$^{-2}$)</th>
<th>Number of grids</th>
<th>NMB$^a$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York City, NY</td>
<td>Mean: 5.90, STD: 3.45</td>
<td>Mean: 4.29, STD: 2.19</td>
<td>64</td>
<td>38%</td>
<td>0.82</td>
</tr>
<tr>
<td>Los Angeles, CA</td>
<td>Mean: 6.42, STD: 3.80</td>
<td>Mean: 5.58, STD: 2.43</td>
<td>48</td>
<td>15%</td>
<td>0.64</td>
</tr>
<tr>
<td>Chicago, IL</td>
<td>Mean: 5.89, STD: 2.03</td>
<td>Mean: 4.33, STD: 1.13</td>
<td>45</td>
<td>36%</td>
<td>0.72</td>
</tr>
<tr>
<td>Houston, TX</td>
<td>Mean: 4.97, STD: 2.09</td>
<td>Mean: 3.36, STD: 1.20</td>
<td>32</td>
<td>48%</td>
<td>0.72</td>
</tr>
<tr>
<td>Phoenix, AZ</td>
<td>Mean: 3.48, STD: 1.47</td>
<td>Mean: 3.24, STD: 0.89</td>
<td>24</td>
<td>7%</td>
<td>0.25</td>
</tr>
<tr>
<td>Philadelphia, PA</td>
<td>Mean: 4.20, STD: 1.34</td>
<td>Mean: 3.34, STD: 0.61</td>
<td>35</td>
<td>26%</td>
<td>0.74</td>
</tr>
<tr>
<td>Atlanta, GA</td>
<td>Mean: 3.23, STD: 1.17</td>
<td>Mean: 2.96, STD: 0.59</td>
<td>50</td>
<td>9%</td>
<td>0.50</td>
</tr>
<tr>
<td>Miami, FL</td>
<td>Mean: 2.90, STD: 0.68</td>
<td>Mean: 2.33, STD: 0.57</td>
<td>24</td>
<td>25%</td>
<td>0.74</td>
</tr>
<tr>
<td>Dallas-Fort Worth, TX</td>
<td>Mean: 4.82, STD: 1.45</td>
<td>Mean: 3.63, STD: 0.77</td>
<td>30</td>
<td>33%</td>
<td>0.73</td>
</tr>
<tr>
<td>Washington, DC/Baltimore</td>
<td>Mean: 3.78, STD: 1.11</td>
<td>Mean: 3.05, STD: 0.47</td>
<td>38</td>
<td>24%</td>
<td>0.83</td>
</tr>
</tbody>
</table>

$^a$NMB: Normalized mean bias. $^b$STD: Standard deviation within the city.
consistent with the fuel-based inventory evaluations based on P-3 aircraft and DOAS measurements during the California Nexus Study in 2010 (Kim et al., 2016). The normalized mean bias of the model relative to TROPOMI NO\textsubscript{2} observations range from 7% to 48% across these cities. Houston shows the largest difference (+48%) in the model compared to TROPOMI. The potential low bias in satellite data due to uncertainties of cloud pressure retrievals can be one of the factors to the discrepancies for cities near oceans with low-lying cloud. It is possible that the emissions are overestimated in the current inventory for Houston, which is a city with high industrial and port-related emissions, and exhibits strong gradients in NO\textsubscript{2} with environmental justice implications (Demetillo et al., 2020). The coefficients of determination of Phoenix (0.25) and Atlanta (0.50) are relatively low compared to other cities, indicating spatial discrepancies between model and satellite data within these cities that could be related to emissions. The retrieval uncertainties from surface albedo can also contribute to these discrepancies. Overall, 7 out of the 10 cities exhibited a high \(R^2\) value (>0.70), illustrating the fidelity of intra-urban variations of model and TROPOMI NO\textsubscript{2}, and relevant for assessing environmental justice and variations in NO\textsubscript{2} exposure across communities (Demetillo et al., 2020).

The weekday-weekend variations over cities can provide additional information on the potential sources of biases (Marr & Harley, 2002). Specifically, heavy-duty diesel truck emissions are known to decrease significantly on weekends versus weekdays (Kim et al., 2016; Marr & Harley, 2002), and this weekend-weekday difference has been observed from space previously (Russell et al., 2010). Here, the decline of NO\textsubscript{2} emissions from heavy-duty diesel trucks during weekends is confirmed both from our bottom-up emissions inventory and from TROPOMI NO\textsubscript{2} columns as shown in Figure 3. During the weekend, the traffic emissions are dominated by light-duty vehicles at 13:00 p.m. local daylight time both in urban and rural areas around the city. On weekdays, the NO\textsubscript{2} emissions and NO\textsubscript{2} columns include significant contributions from heavy-duty diesel trucks (Marr and Harley, 2002). As demonstrated by Figure S10 in Supporting Information S1, the model biases decrease with increasing fractions of the on-road. This suggests that on-road emissions may be better estimated compared to other more uncertain sources (off-road, areawide) in urban areas. Taking Los Angeles (LA) as an example, the increase of the model bias in LA on weekends compared to that on weekdays (11% vs. 35%) suggests that there is potentially an overestimate in emissions from a source that does not vary greatly between weekdays and weekends, such as passenger vehicles (McDonald et al., 2014). Since on-road transportation emissions account for ~70% of the anthropogenic NO\textsubscript{2} emissions in Los Angeles (Table S2 in Supporting Information S1), it is possible that light-duty gasoline vehicle emissions are slightly overestimated in FIVE for California. A recent analysis by Yu et al. (2021) shows that EMFAC reports light-duty gasoline NO\textsubscript{2} emissions that are lower than FIVE, albeit within the uncertainty limits of FIVE (Yu et al., 2021). California implements tailpipe emission standards that can precede and differ from other regions of the country (Council, 2006). There could also be uncertainties in other sources of NO\textsubscript{2} emissions such as from ports, railyards, and point sources that require a more detailed investigation. Nationally, the mix of light-duty and heavy-duty vehicle NO\textsubscript{2} emissions, along with point and area source emissions, simulated in WRF-Chem appear to capture weekend-weekday modulation in NO\textsubscript{2} columns. The WRF-Chem model reflects changes in weekday/weekend emissions input into the model, as well as resulting impacts on NO\textsubscript{2} lifetime and chemistry which are explicitly accounted for in a CTM (Valin et al., 2014). At a spatial resolution of 12 × 12 km\textsuperscript{2}, the biases associated with model resolution on NO\textsubscript{2} lifetimes is expected to be minimal for urban areas (Valin et al., 2011).

### 3.4. Model Evaluations Over Background and Rural Regions

As illustrated in Figure 2e, WRF-Chem agrees well with TROPOMI (slope = 0.94, NMB = −3%), and 21% lower than OMI observed NO\textsubscript{2} data over background and rural regions. We grouped the “background and rural” grid cells that are located far away from point sources or cities (>50 km). Like for point sources and large cities, the agreement between the model and satellite tropospheric NO\textsubscript{2} is within the uncertainties of both. The model-satellite differences of up to 20% is expected and consistent with previous studies (Silvern et al., 2019). The better correlation coefficient of TROPOMI versus OMI (\(R^2 = 0.78\) vs. 0.68) over relatively clean area benefits from the higher spatial resolution, and more valid data (17,951 vs. 15,042) in TROPOMI.

Steady decreases of NO\textsubscript{2} emissions from transportation and power generation over urban areas suggest that the contributions from other sectors (areawide, soil, O&G production, fire, lightning, etc.) are becoming increasingly important and should be accounted for carefully (Silvern et al., 2019). To address this issue, we conducted a series of sensitivity model simulations for July 15 to August 14 to identify the regions with various dominant sources: soil, oil & gas production, wildfires, and lightning. The details of sensitivity model setup are summarized in
Table 1. We used the “zero-out” method to estimate the emission contributions to tropospheric NO$_2$ columns for each source, as shown in Figure 4. For regions with sectoral contribution >20%, we show the histograms of WRF-Chem and TROPOMI NO$_2$ columns (see right columns of Figure 4). We did not include wildfire or lightning NO$_x$ emissions in our baseline runs, but illustrate their importance in sensitivity test runs as follows, considering the high uncertainties involved, and the minor impact on tropospheric NO$_2$ columns over CONUS on cloud-free pixels in pre-COVID period, which are detailed and illustrated below.

Figure 3. Weekday-weekend variations of NO$_2$ tropospheric columns for the top 10 cities by population in WRF-Chem (red) and TROPOMI (blue). The whiskers denote the standard deviations of the NO$_2$ columns within the city. Gray dots show the normalized mean biases of WRF-Chem versus the TROPOMI NO$_2$ trop. columns.
Figure 4. Percent contributions of emissions from non-point and non-urban sources to CONUS tropospheric NO$_2$ columns for (a) Soil, (b) Oil and gas production, (c) Wildfire, and (d) Lightning at TROPOMI passing time (13:00–14:00 p.m). Results are based on the sensitivity simulations for July 15 to August 14 in 2018. The histogram figures in the right panels are for the regions where the column contributions of NO$_x$ source is greater than 20%.
3.4.1. Soil

BEIS3 calculates the soil NO\textsubscript{2} emissions online with meteorology in WRF-Chem as 4541 Mg d\textsuperscript{-1} (expressed as metric tons of NO\textsubscript{2} per day) during the sensitivity study period, and contributes 11% of the domain-wide emissions total (Table 2). These estimates (0.04 Tg N\textsuperscript{-1} month\textsuperscript{-1}) are consistent with a previous bottom-up inventory of NO\textsubscript{2} from soil (−0.06 Tg N\textsuperscript{-1} month\textsuperscript{-1}) (Hudman et al., 2012) in July 2005. It has been shown that soil NO\textsubscript{2} inventories (Hudman et al., 2012) have overestimated 40% over Central US through top-down inversions of OMI observations (Vinken et al., 2014a). Airborne measurements demonstrate that agricultural soil NO\textsubscript{2} emissions have been largely overlooked in California, which would potentially increase the NO\textsubscript{2} budget by 20%–51% (Almaraz et al., 2018). Recent modeling work based on an updated soil scheme estimates that soil contributes 40.1% of the total NO\textsubscript{2} emissions in July 2018 over California (Sha et al., 2021). The above top-down evaluations indicate that uncertainties are still high for the soil NO\textsubscript{2} emissions regionally.

Figure 4a demonstrates that the highest sensitivities of soil emissions to the tropospheric NO\textsubscript{2} column are over the central US and the Central Valley of California with relative contributions larger than 20%. The modeled NO\textsubscript{2} columns over the central US grid cells are dominated by those emissions from agricultural activities (with contributions >40%). Over these two regions, WRF-Chem agrees with the TROPOMI observed NO\textsubscript{2} columns (see Figure 4). Excluding the soil emissions worsens the bias from −7% to −15% over CONUS, and from −10% to −27% over the soil dominant regions. If the model-satellite discrepancies are solely due to soil NO\textsubscript{2} emissions over background regions, we estimate BEIS could underestimate soil emissions by 8% (TROPOMI) to 57% (OMI). The differences between TROPOMI and OMI tropospheric columns can be presumably due to the different cloud retrievals and AMFs (Wang et al., 2020). Especially over the background and rural regions, TROPOMI data tend to be 20% lower than those of OMI (as shown in Figure 2). OMI retrievals are biased high in remote areas within 20% differences (Lamsal et al., 2014), and this can partly explain the discrepancies. More independent validations of TROPOMI NO\textsubscript{2} retrievals over soil-dominant regions are needed in future work.

3.4.2. Oil and Gas Production

Because of a large increase in oil and gas production activities between 2017 and 2019, an increasing trend of NO\textsubscript{2} columns over O&G basins is visible from space (Dix et al., 2020). In this work, we included the FOG Inventory with up-to-date oil and gas NO\textsubscript{2} and VOC emissions over CONUS in our base case modeling (Francoeur et al., 2021; Gorchov Negron et al., 2018). NO\textsubscript{2} emissions from drilling, oil and gas production, and flaring provide an additional source of background NO\textsubscript{2} than agricultural soils over the central US. WRF-Chem shows good correlations with TROPOMI over O&G basins as shown in Figure 2f ($R^2 = 0.78$, slope = 1.02, NMB = +4%). Excluding oil and gas emissions increases the model bias from +5% to −20% in the major O&G production basins (see Figure 4b), including the Permian, Eagle Ford, Haynesville, Barnett, Bakken over the central US, and Marcellus over the eastern US, illustrating the importance of O&G sources in certain regions of CONUS.

3.4.3. Wildfires

During the summer of 2018, there were several wildfires affecting the air quality over the western US (Lindaas et al., 2021). Wildfire emission inventories are highly uncertain (>100%) when applied in regional air quality models, due to the difficulties in quantifying the burned area, the fuel and fire conditions, and the height of the plumes (French et al., 2004; Urbanski et al., 2011). We conducted a sensitivity test with the 3BEM (Longo et al., 2010) fire emission estimates during wildfire season (July 15 to August 14) in 2018. As shown in Figure 4c, wildfire emissions lead to NO\textsubscript{2} column increases of up to 30% over the western US, reducing the bias from −7% to −1% over the CONUS, and from −17% to +13% over wildfire dominant areas. The wildfire emissions’ impact on the background tropospheric NO\textsubscript{2} columns is comparable with those from soils, although located in different regions (Western US vs. Central US). Wildfires also impact the accuracy of the TROPOMI NO\textsubscript{2} data due to high aerosol loads, which interfere with the satellite trace gas retrieval.

3.4.4. Lightning

The inclusion of lightning NO\textsubscript{2} in the free troposphere can substantially contribute to the NO\textsubscript{2} vertical profile and AMF, and thus satellite retrievals (Allen et al., 2021; Pickering et al., 2016; Zhu et al., 2019). Currently, there is a range of an order of magnitude in the lightning NO\textsubscript{2} emissions reported in the literature (Nault et al., 2017; Schumann & Huntrieser, 2007; Zhu et al., 2019), because of the dependence of lightning NO\textsubscript{2} on lightning flash rates and potentially other factors like the length of the flash (Pollack et al., 2016) which depend on storm parameters.
and their uncertainties (Barth et al., 2019). Aircraft measurements along with modeling studies have shown that lightning is the major source of upper tropospheric \( \text{NO}_x \) columns over the eastern US (Bertram et al., 2007; Cooper et al., 2006, 2007, 2009; Lapierre et al., 2020; Nault et al., 2017; Zhu et al., 2019). The average contribution to \( \text{NO}_x \) columns is expected to be 19% on average over the Southeast US although with large variations (Zhu et al., 2019). In this study, we applied one of the default lightning schemes based on cloud top height in WRF-Chem in the sensitivity test (Wong et al., 2013), as described in Section 2.1. Lightning contributes \( \sim 6\% \) of the total \( \text{NO}_x \) emissions over the whole domain, and exceeds 20% of the tropospheric \( \text{NO}_x \) columns at a regional scale, especially over Southeast US, where mid-day convective storms are prevalent, and southern Arizona, where outflow from the Sierra Madre storms likely occur (see Figure 4d and Figure S2 in Supporting Information S1).

In this case, lightning \( \text{NO}_x \) appears to have limited impacts on model-satellite comparisons over CONUS with lightning turned on (\( \text{NMB} = -4\% \)) and off (\( \text{NMB} = -7\% \)) when applying the high filtering factor (\( qf > 0.75 \)) in satellite data resampling. Without performing a detailed evaluation of flash rate prediction or additional sensitivity simulations of the lightning and lightning-\( \text{NO}_x \) parameterization, there is some uncertainty in these results. For example, Wong et al. (2013) and Zhu et al. (2019) reduced the flash rate prediction by a factor of 9 and 2, respectively, for their simulations using cloud top height to predict flash rate. In addition, Zhu et al. (2019) used a much higher \( \text{NO}_x \) production per flash (500 moles/flash) than we used here (125 moles/flash).

Apart from the above sources, full aviation emissions can also affect the modeled \( \text{NO}_x \) vertical profiles, AMFs, and the satellite retrievals. In this study, we did not apply aviation emissions for cruising in model simulations because they are not included in the default WRF-Chem package, and not updated on a regular basis like NEI/FIVE. Only take-off and landing emissions are included here. The contribution of \( \text{NO}_x \) emissions from aviation to the total emissions in the United States is reported to be less than 1% in 2005 (FAA, 2005). But in recent years, aviation emissions are expected to have increased due to increased flight activities, while the trends of mobile source \( \text{NO}_x \) emissions at the surface have been declining (McDonald et al., 2012; Yu et al., 2021). The model-ready aviation emission inventory with resolved vertical distributions are outdated and should be updated in the near future (Simone et al., 2013; Stettler et al., 2011, 2013). Additionally, satellite retrievals are more sensitive to \( \text{NO}_x \) released at high altitude, like the emissions from aircraft. Better correlations between model and satellite observations are expected especially over remote regions where \( -3 \sim -21\% \) biases are found in our assessment.

4. Conclusions

In this study, we applied the high-resolution \( \text{NO}_x \) column retrievals from TROPOMI to evaluate WRF-Chem using an updated fuel-based emission inventory over the CONUS for summer 2018. We updated fuel-based emission inventories to 2018, including FIVE for mobile sources, FOG for oil & gas production, and combined the fuel-based inventory with CEMS data for point sources, CAMS for Canada/Mexico, and NEI17 for other point and areawide sources. We find that the FIVE18 and NEI17 are in good agreement in contrast to discrepancies reported between FIVE and the NEI in the past. This study then performs a sector-by-sector analysis of the newly developed inventory to 2018, including FIVE for mobile sources, FOG for oil & gas production, and combined the inventory with TROPOMI and OMI \( \text{NO}_x \) data. The \( \text{NO}_x \) vertical profiles in the model are used to revise the AMF in the standard \( \text{NO}_x \) product to facilitate a direct comparison. Applying the updated AMF based on our regional CTM increased the tropospheric \( \text{NO}_x \) columns by 20% over urban areas, and decreases the columns by 3% on average over background and rural regions. Revising the AMF with a regional CTM reduces half of the low bias of TROPOMI tropospheric \( \text{NO}_x \) (\( \sim 50\% \)) reported previously when compared to surface-based remote sensing.

Tropospheric \( \text{NO}_x \) columns of WRF-Chem and TROPOMI (OMI) show good spatial and temporal correlations over the whole domain with correlation coefficients up to 0.78 (0.69) and a normalized mean bias of \( -3\% \) (\( -21\% \)). WRF-Chem reproduces tropospheric \( \text{NO}_x \) columns over isolated point sources plus urban regions within \( -6\% \) (OMI) to \( +8\% \) (TROPOMI), but tends to under-predicts tropospheric \( \text{NO}_x \) columns over background and rural regions by \( -21\% \) (OMI) to \( -3\% \) (TROPOMI). Strong agreement between model and satellite data is found for point sources (\( \pm 10\% \) differences), which is expected because of the direct stack monitoring data (CEMS) used to develop the power plant emissions. The modeled columns are 27% higher than those of TROPOMI, and 4% higher than the OMI data for the top 10 most populated cities. These differences can be largely attributed to the low biases of TROPOMI (\( -20\% \)) over polluted regions, which have been assessed comprehensively based on ground measurements in previous work. The decline of \( \text{NO}_x \) emissions from heavy-duty diesel trucks during weekends is confirmed both from our bottom-up emissions inventory and from TROPOMI \( \text{NO}_x \) columns.
We found a good agreement between model and TROPOMI over the background and rural regions across the CONUS ($R^2 = 0.78$, slope = 0.94, NMB = $-3\%$), and 21% model low bias compared to OMI ($R^2 = 0.68$, slope = 0.78, NMB = $-21\%$). Acknowledging the potential satellite bias for low concentrations which can be related to the stratosphere-troposphere separation, more validations over background regions is needed in future work. Further sensitivity analyses were conducted to identify the dominant source for each grid and the effect of emission updates on NO$_2$ columns at continental and regional scales. The results illustrate the important contributions of tropospheric NO$_2$ columns from various non-point and non-urban sources, including soils (11.7% average over CONUS), oil and gas production (4.1%), wildfires (10.6%), and lightning (2.3%). Regionally, over the central US, soils can be the major contributing source (>20%), with additional contribution from oil and gas production activities. Soils and wildfires are important contributing sources over the rural and background regions in the western US. Lightning NO$_x$ can be one of the major contributors to the model bias over Southeast US and southern Arizona with large regional variations. Note these sensitivity results are limited to summertime, and source contributions may change by season across the country. Acknowledging the ±20% uncertainty in satellite tropospheric NO$_2$ retrievals as inferred from the comparisons of TROPOMI and OMI, as well as satellite validation results based on ground-based measurements, overall we do not find significant discrepancies of online simulations of soil and lightning NO$_x$ between WRF-Chem and satellite data.

Another potentially important NO$_x$ emission source currently not included in the model is domestic and international aviation at cruising altitude in the free troposphere; only takeoff and landing emissions are included here. Future work is needed to quantify the impact of updated aviation emissions on the satellite retrievals and model simulations over urban, as well as background and rural regions. Future development of FIVE may consider the inclusion of aviation along with other mobile source engine categories. This work also provides an important baseline of a typical year against which sharp changes in anthropogenic NO$_x$ emissions due to the COVID-19 pandemic can be assessed.

**Data Availability Statement**

The TROPOMI L2 NO$_2$ swath data are downloaded from the Sentinel-5P Pre-Operations Data Hub (https://s5phub.copernicus.eu/dhus/#/home). The OMI tropospheric NO$_2$ column V003 Level 2 swath data were obtained from the NASA’s Earth Observing System Data and Information System (EOSDIS). The PANDORA spectrometer observations are accessed from the NASA Pandora project (https://pandora.gsfc.nasa.gov/). The 2018 anthropogenic emissions files are publicly archived at the NOAA Chemical Sciences Laboratory (CSL) COVID Air Quality Study website: https://csl.noaa.gov/groups/csl7/measurements/2020covid-aqs/emissions/.

**Acknowledgments**

This study was supported by NASA ROSES ACMP (80NSSC19K0979 and 80NSSC19K0950) and Cooperative Agreement with CIRES, NA17OAR4320101. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect the views of NOAA or the Department of Commerce. The authors also thank NOAA’s High Performance Computing system to support the WRF-Chem modeling.

**References**


Liu, F., Beirle, S., Zhang, Q., Dörner, S., He, K., & Wagner, T. (2016). NOx lifetimes and emissions of cities and power plants in polluted background estimated by satellite observations. *Atmospheric Chemistry and Physics, 16*(8), 5283–5298. https://doi.org/10.5194/acp-16-5283-2016


